Consistencies and Inconsistencies in Personality Ratings Among Knowledgeable Other Reporters

Megan E. Gillespie

Appalachian State University

Abstract

Most personality assessments rely on self-report, but the Realistic Accuracy Model (RAM) proposes that personality may be accurately perceived by others. The strongest reporters of personality besides the targets themselves are knowledgeable others. Research supports that spouses exhibit the strongest correspondence between self-reports and other-reports, followed by family members, then friends. The Self-Other Knowledge Asymmetry model proposes that traits with high visibility, such as extraversion and openness to experience, are more accurately perceived by others than are traits with low visibility, such as neuroticism. However, both self-reports and other-reports may be vulnerable to biases. This study uses a sample of 197 targets, 197 friends, and 151 family members to examine the consistencies and inconsistences of personality ratings across nine traits. Based on the literature, I hypothesized that self-family reports and self-friend reports would be positively correlated, and that there would be stronger correlations between self-family reports than between self-friend reports. I also hypothesized that there would be a positivity bias for family reports of the target's personality, and no significant bias among friend reports the target's personality, relative to the self-reports. Both sets of knowledgeable others (friend-reporters and family-reporters) were consistent with their personality judgments of the target, with similar average correlations. Parents tended to be positively biased, relative to self-reports, for the traits Extraversion, Agreeableness, Self-Esteem, and Integrity. Friends tended to be negatively biased, relative to self-reports, for the traits Openness to Experience and Intellect.

Consistencies and Inconsistencies in Personality Ratings Among Knowledgeable Other Reporters

There has been much research in the past few decades examining the accuracy of personality judgments. The Five Factor Model of personality is the most widely accepted factor analytical structure used by researchers (Goldberg, 1990). Within the model, personality is broken down into five main domains: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness. Extraversion, also referred to as surgency, can be described as how outgoing or energetic a person is. Agreeableness can be described as how friendly, compassionate, or easy-going a person is. Conscientiousness is the trait that includes efficiency, dependability, and organization. Neuroticism, also referred to as emotional stability, is the tendency to experience anxiety, anger, and depression. Openness, also referred to as culture or intellect, generally reflects the appreciation one has for adventure, curiosity, and art.

Several tools can be used to measure personality, one of which is the Revised NEO Personality Inventory (NEO-PI-R; Costa & McCrae, 1992). The NEO-PI-R is a self-report personality instrument that researchers use to get personality information directly from their subjects. Self-reported assessments of personality are most commonly used in personality research, however, reports from others are also sometimes utilized (Funder & Colvin, 1988). Knowledgeable others, family members, friends, and sometimes strangers are some of the other types of informants used to get personality reports.

Research has shown that the most accurate reporter of a target's personality is the target themselves. However, knowledgeable others have also proven to be good reporters of personality traits. The average of self-other correlations is .49 (Allik, de Vries, & Realo, 2016). According to Connelly and Ones (2010), the accuracy of personality judgments does not increase with the

quantity of observations, but it does increase with the quality of the information available to the observer. This would indicate that the quality of the relationship between the target and the knowledgeable other would affect the accuracy of personality judgments. Therefore, self-family correlations should be stronger than self-acquaintance correlations.

Self-spouse (or significant other) relationships exhibit some of the strongest correlations for personality judgments (Connelly & Ones, 2010; Kurtz, Tarquini, & Iobst, 2008). In a meta-analysis completed by Connelly and Ones (2010), correlations ranged from .5 to .6 between spouses and targets ratings, and ranged from .5 to .7 for significant others and target ratings. Similar to the results found by Connelly and Ones (2010), Allik et al. (2016) found self-spouse correlations averaging .5. However, some research has demonstrated that self-spouse correlations decrease when adjusted for social desirability responding (Kurtz et al, 2008). This suggests that spouses may have different motivations for rating personality traits than other knowledgeable others. It also suggests that there may be shared biases between the target and the spouse.

Following self-spouse correlations in strength are self-family correlations. Self-parent and self-sibling correlations are of comparable strength to one another (Allik et al., 2016; Connelly & Ones, 2010). Self-parent correlations have an average of .3 to .4 on all traits, while self-sibling correlations have an average of .5 (Connelly & Ones, 2010). However, the slight differences in these correlations are congruent with the theory that the quality of the relationship will strengthen the correlations of the raters. It can be expected that self-sibling correlations will be greater than self-parent correlations because siblings are able to see each other in more environments than parents would (Connelly & Ones, 2010). Parents also have a dominant structured relationship with the target, which could affect their accuracy in ratings (Connelly & Ones, 2010).

Self-friend correlations are almost as diverse as self-family correlations. Surprisingly, close friends/acquaintances had stronger correlations with the target for most traits than best friends (Connelly & Ones, 2010). However, roommates were shown to have the second strongest correlations with the target, after target-dating partner correlations, with correlations ranging from .4-.6 for all traits.

There have been very few studies done to examine the potential bias of self-ratings of personality, compared to family-ratings and friend-ratings. One study in particular (Van Heck & Dijkstra, 1985) has suggested that knowledgeable others tend to make more general heuristic judgments about the targets, than the targets make about themselves, due to the limited number of settings the knowledgeable others see the targets in (Van Heck & Dijkstra, 1985). Also affecting their ratings are the types of settings that knowledgeable others associate the target with (Van Heck & Dijkstra, 1985). This observability limitation may bias the perceptions of knowledgeable others, although different reporters may be differentially vulnerable to bias. For example, a recent study has suggested that parents may have stronger biases about a target than friends do (Kurtz, Tarquini, & Iobst, 2008).

As important as the quality of the relationship is between the target and its rater, there are also other factors that contribute to making an accurate personality judgment. The Realistic Accuracy Model (RAM) outlines four conditions that must be satisfied before an accurate judgement can be made (Funder, 1995). The first condition is that the target must do something to demonstrate the trait being judged. The second condition is that the rater must have access to the demonstrated behavior of the trait. The third condition is that rater must register the trait being demonstrated, and the last condition is that the rater must utilize the information they gather from the demonstration to make an accurate judgment. Therefore, in order to make the

best judgment, there must be a good judge, judging a strong trait, using good information (Funder, 1995).

Another model used for making personality judgments is the Self-Other Knowledge Asymmetry (SOKA) model (Vazire, 2010). This model is used to predict the accuracy of self-other judgments. It takes into consideration the effects that trait observability and evaluativeness can have on a judgment's accuracy. Observability refers to how accessible the behavioral indicators of the trait may be to an observer, whereas evaluativeness refers to how positively or negatively valenced a trait is considered to be. This model also examines the role of the knowledgeable other who is making the judgment.

SOKA (Vazire, 2010) suggests that some traits are more easily judged than others. Traits with higher visibility, such as extraversion, are easier to pick up on than traits with low visibility, such as neuroticism. It has also been suggested that although targets have better accuracy than knowledgeable others at reporting internal traits such as neuroticism, knowledgeable others generally have better accuracy judging external traits such as extraversion or openness. According to Vazire and Carlson (2011), targets have better accuracy judging traits with low observability and low evaluativeness, such as anxiety. Friends of the target have better accuracy judging traits high observability and low evaluativeness, such as oral communication, and traits with low observability and high evaluativeness, such as intelligence.

The Present Study

One purpose of the present study is to study the consistencies and inconsistencies of personality ratings among self-reports of personality and two sets of knowledgeable-other reports. The two types of knowledgeable others used in this study are family members and friends. Consistencies will be examined by calculating correlations between self-friend and self-

family reports. In addition to Five Factors Personality ratings, there will also be calculations of the correlations for self-esteem, integrity, intellect, and attractiveness.

Based on the literature, I hypothesize (Hypothesis 1) that self-family reports will be positively correlated. In addition to self-family reports, I hypothesize (Hypothesis 2) that self-friend reports will be positively correlated as well. I also expect stronger correlations between self-reports and family members, than between self-reports and friends (Hypothesis 3). This hypothesis is based on the theory that the quality of the relationships affects the strength of the correlation. I expect that stronger correlations among traits such as extraversion and openness, because of their greater visibility.

I also hypothesize (Hypothesis 4) that there will be positive bias for family-reports of personality, relative to self-reports and that there will be no significant bias among friend-reporters (Hypothesis 5). Although there is little research concerning the biases of personality judgments, this hypothesis is based on the theory that knowledgeable others have different types of motivations when rating personality traits. These biases will be examined through mean differences among the three reporters.

Methods

Participants

The targets (121 women, 76 men) used in the current study were undergraduate students recruited from a university in North Carolina. In exchange for their participation, the students were given credit for a research requirement for a related psychology course. On the day of the study, each target was expected to bring a close friend and email contact information of a parent or close family member. The friend raters were also undergraduate college students (116 women, 74 men, 7 not reported). The majority of the family sample consisted of parents, specifically

mothers. The final sample consisted of data from 197 self-reports, 197 friends, and 150 family members.

Measures

The tool used to measure the target's personality was the International Personality Item Pool (IPIP; Goldberg, 1999) version of the revised NEO Personality Inventory (NEO PI-R; Costa & McCrae, 1992). The IPIP scale consists of 10 items for each of the 5 personality constructs: openness to experience (α =.82), conscientiousness (α =.81), extraversion (α =.86), agreeableness (α =.77), and neuroticism (α =.86). The response format for each item was a 5-point rating scale, ranging from *(1) very inaccurate* to *(5) very accurate*. The knowledgeable others were presented with a slightly alternate version of these scales, with the questions referring to the target and not the actual knowledgeable other completing the scale.

The tool used to measure intelligence was an IPIP scale of intellect that is similar to Gough's California Psychological Inventory (CPI; Johnson, 2000). The IPIP scale consists of 10 items and uses a 5-point rating scale, ranging from *(1) very inaccurate* to *(5) very accurate,* with a reliability of .85. Additional scales were used to measure the traits of self-esteem, integrity, and attractiveness. Self-esteem was measured using Rosenberg's IPIP scale of self-esteem (1965). The IPIP scale consists of 10 items, and has a reliability of .84. Integrity was measured using Peterson and Seligman's (2004) IPIP scale of integrity. The IPIP scale consists of 9 items, with a reliability of .72. Attractiveness was measured using Saucier's (1997) IPIP scale of attractiveness. The scale consists of 10 items and uses a 5-point rating scale, ranging from *(1) very inaccurate* to *(5) very accurate,* with a reliability of .66.

Procedure

On the day of the study, targets completed self-reported measures of the instrument. Friends were given versions of the instrument that had been modified to be used as other-reports and were asked to rate the target on various measures. Family members were emailed the modified other-report versions of the instrument and were also asked to rate the target on various measures.

Analytic Plan

For Hypothesis 1 and 2, the direction and strength of the correlations will be examined. For Hypothesis 3, the difference in correlations will be examined using an *r*-to-*z* transformation. For Hypotheses 4 and 5, the possible biases will be examined using a dependent *t*-test.

Results

Table 1 shows the direction and the strength of the correlations examined for Hypotheses 1 and 2, as well as the correlations between family-reports and friend-reports. I hypothesized that there would be positive correlations between self-reports and family-reports of personality traits. According to the data, there were statistically significant positive correlations between self-reports and family-reports for all nine examined traits. The mean correlation between self-reports and family-reports was .45. Therefore, Hypothesis 1 was supported by the data.

I also hypothesized that there would be positive correlations between self-reports and friend-reports of personality traits. According to the data, there were statistically significant positive correlation for all nine traits. The mean correlation between self-reports and friend-reports was .45. Therefore, the hypothesis was supported by the data.

Table 2 presents results from Hypothesis 3 that the self-family correlations would be stronger than the self-friend correlations. According to the data, there were no significant

differences between the correlations for all nine traits. These results indicate that the hypothesis was not supported by the data.

Tables 3 and 4, and Figure 5 present results from Hypotheses 4 and 5 by examining the means and *t*-tests among the three reporters. For Hypothesis 4, I hypothesized that there would be a positive bias for family-reports of personality compared to self-reports. According to the data, there were significant mean differences for four of the nine traits: extraversion, $t(150) = -3.36$, $p = .001$; agreeableness, $t(150) = -5.09$, $p < .001$; self-esteem, $t(150) = -2.483$, $p = .014$; and integrity, $t(150) = -2.996$, $p = .003$, such that the parents rated the targets more positively than the self-reports. These negative mean differences indicate a positive bias for these four traits, suggesting that the hypothesis was partially supported by the data.

For Hypothesis 5, I hypothesized that there would be no bias among friend-reports of personality compared to self-reports. According to the data, there were significant mean differences for two of the nine traits: openness to experience, $t(197) = 4.07$, $p < .001$, and intellect, $t(197) = 3.950$, $p < .001$, such that the friends rated the targets more negatively than the self-reports. These positive mean differences indicate a negative bias for these two traits, suggesting that the hypothesis was not supported by the data.

Discussion

One purpose of the current study was to examine the consistencies among the different types of reporters, specifically self, family, and friend reporters. I expected there to be positive correlations among all three types of reporters. Previous studies have supported this by repeatedly showing significant correlations among self-reports and knowledgeable other-reports (Allik et al., 2013). The current study also supported this hypothesis by showing significant correlations among all three reporters for all nine traits. As demonstrated by the results in Table

1, some traits showed stronger correlations among the reporters than other traits. For example, Extraversion and Openness to Experience had the strongest correlations out of all the traits judged. This supports Vazire and Carlson's (2011) theory that traits with high observability are more easily judged than traits with low observability, such as neuroticism.

In addition to the expectation of positive correlations among the three reporters, I also expected there to be stronger correlations among self-family reports than the self-friend reports. The rationale behind this expectation was that the strength of the correlations is increased as the quality of the relationship increases (Connolly & Ones, 2010; Vazire, 2010). However, the current study showed no significant difference among the correlations of self-family reports and self-friend reports for any of the examined traits. An interpretation of these results could be that the participants in this study had approximately the same quality of relationship with both types of reporters.

After looking at the consistencies between knowledgeable others judgments of personality, I examined the inconsistencies between reporters. By looking at the mean differences among the reporters, I was able to determine if the knowledgeable others rated the participants more positively or negatively than the participants rated themselves for each of the nine traits. Figure 1 provides a useful illustration of the differences between the raters' judgments for all nine traits, along with the direction of the differences.

For the traits of extraversion, agreeableness, self-esteem, and integrity, parents were positively biased in their answers compared to the participants. In other words, parents were more likely to rate their children as having higher levels of extraversion, agreeableness, self-esteem, and integrity than their children rated themselves. These results were congruent with the study predictions.

Very few studies have examined the biases among knowledgeable other reports, therefore making it difficult to compare and interpret the results with other studies. However, most researchers agree that different raters have different motivations when making personality judgments. For example, parents may feel more inclined to rate their children higher on personality traits than other raters because their child is representative of the rater. Parents can also be more subjective when rating their children than other raters may be.

In contrast to family raters, friend raters were more likely to rate the participants more negatively for the traits of openness to experience and intellect than the participants and family raters, resulting in a negative bias. While friends can still be subjective in their ratings, they may be more objective in their ratings than parents because the parents' self interest in their offspring may increase their motivation to see their children more positively. Another explanation for these results is the type of settings that friends are more likely to see the participants in compared to parents. Because both the participants and the friend raters in this study were undergraduate students living in a college setting, this could explain why the participants were rated as having lower levels of intellect and openness to experience, each of which has aspects of intelligence in its construction. Participants are also more likely to have an informal relationship with their friends than with their parents, which could also explain the negative biases.

One limitation of this study was that all the participants attending the same university. Therefore, there was little variety in the types of participants being selected. These participants were all relatively from the same area, making them unrepresentative of the population. These participants were also all college aged students, limiting the age range of people this study could apply to. Another limitation is the accuracy of the ratings. The bias in the current study was examined relative to self-report, not to an absolute criterion. Since target ratings are considered

to be the most accurate raters of themselves, they were used as the criterion variable when examining biases in this study. However, there are still elements of bias in their self-judgments as well, making it difficult to measure true bias among reporters. Using behavioral reports or observations in future research might better inform this issue.

In the future, I would like to expand bias research to all types of knowledgeable-other reporters, including spouses, significant others, co-workers, roommates, and other relationships. I would also like to separate the family category into multiple categories, including mother, father, brother, sister, and grand-parents ratings. In addition to these extensions of the current study, I would also be curious if these results would stay static for all ages, or if they would change.

References

Allik, J., de Vries, R. E., & Realo, A. (2016). Why are moderators of self-other agreement difficult to establish? *Journal of Research in Personality, 63,* 72-83.

Connelly, B. S., & Ones, D. S. (2010). An other perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin, 136,* 1092-1122.

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (Neo-PI-R) and NEO Five-Factor Inventory (NEO-FFI): Professional manual*. Odessa, FL: Psychological Assessment Resources.

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review, 102,* 652-670.

Funder, D. C, & Colvin, C. R. (1988). Friends and strangers: Acquaintanceship, agreement, and the accuracy of personality judgment. *Journal of Personality and Social Psychology, 55,* 149-158.

Goldberg, L. R. (1990). An alternative “description of personality”: The Big-Five factor structure. *Journal of Personality and Social Psychology, 59,* 1216-1229.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality measuring the lower-level facets of several five-factor models. In I. Merviedelde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe* (Vol. 7, pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.

Johnson, J. A. (2000). Predicting observers’ ratings of the Big Five from the CPI, HPI, and NEOPI-R: A comparative validity study. *European Journal of Personality, 14,* 1-19.

Kurtz, J. E., Tarquini, S. J., & Iobst, E. A. (2008). Socially desirable responding in personality assessment: Still more substance than style. *Personality and Individual Differences, 45,* 22-27.

Peterson, C., & Seligman, M. (2004). *Character strengths and virtues: A handbook and classification*. New York, NY: Oxford University Press/Washington, DC: American Psychological Association.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University.

Saucier, G. (1997). Effects of variable selection on the factor structure of person descriptors. *Journal of Personality and Social Psychology, 73,* 1296-1312.

Van Heck, G. L., & Dijkstra, P. (1985). The scope and generality of self-other asymmetry in person perception. *European Journal of Social Psychology, 15,* 125-145

Vazire, S. (2010). Who knows what about a person? The Self-Other Knowledge Asymmetry (SOKA) model. *Journal of Personality and Social Psychology, 28,* 281-300.

Vazire, S., & Carlson, E. N. (2011). Others sometimes know us better than we know ourselves. *Current Directions in Psychological Science, 20,* 104-108.

Table 1

Correlations of IPIP Scales among Self and Knowledgable Other Reports

IPIP Scale	Self-Family	Self-Friend	Family-Friend
Extraversion	.57	.59	.50
Agreeableness	.35	.47	.24
Conscientiousness	.40	.49	.29
Neuroticism	.46	.44	.26
Openess	.60	.56	.43
Intellect	.52	.48	.38
Attractiveness	.47	.35	.25
Self-Esteem	.34	.36	.13
Integrity	.32	.31	.29

Note. All of the following are considered statistically significant at $p < .05$ except for the Family-Friend correlation for Self-Esteem ($p = .056$).

Table 2

r to z Transformations between Self-Family Correlations and Self-Friend Correlations

IPIP Scale	Family-Friend *z*	Family-Friend *p* (one-tailed)
Extraversion	0.28	.390
Agreeableness	1.33	.092
Conscientiousness	0.92	.179
Neuroticism	-0.23	.409
Openess	-0.55	.291
Intellect	-0.49	.312
Attractiveness	-1.30	.092
Self-Esteem	0.21	.417
Integrity	-0.10	.460

Note. All differences were nonsignificant at $p > .05$.

Table 3

Means and Standard Deviations for Self and Knowledgable Other Reports

Traits	Self		Family		Friend	
	M	SD	M	SD	M	SD
Extraversion	3.71	0.61	3.87	0.63	3.78	0.69
Agreeableness	3.76	0.59	4.06	0.67	3.77	0.59
Conscientiousness	3.65	0.62	3.69	0.79	3.69	0.71
Neuroticism	2.37	0.68	2.24	0.78	2.33	0.74
Openess	3.61	0.67	3.66	0.67	3.43	0.67
Intellect	3.62	0.69	3.73	0.71	3.42	0.71
Attractiveness	4.00	0.53	4.02	0.57	3.98	0.54
Self-Esteem	3.94	0.57	4.08	0.61	3.93	0.62
Integrity	4.13	0.43	4.30	0.64	4.19	0.52

Note. All variables were measured on a 5-point scale.

Table 4

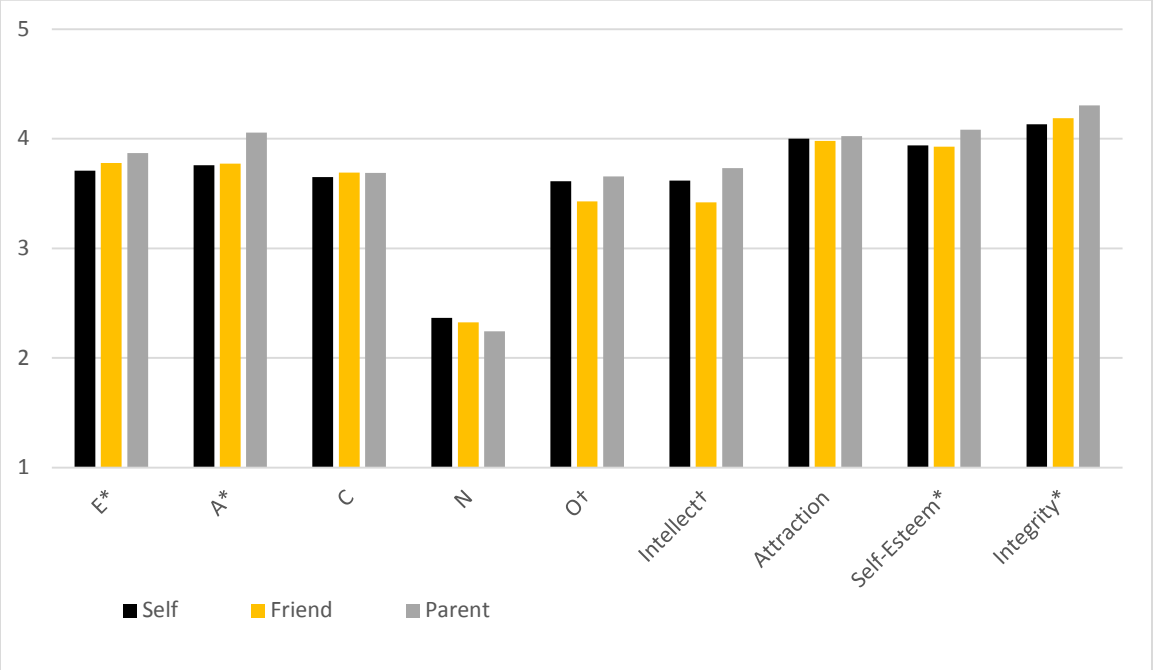
Mean Comparisons of Traits for Knowledgeable Other Reports

Traits	Self-Family *t*	Self-Friend *t*
Extraversion	-3.36*	-1.71
Agreeableness	-5.09*	-0.35
Conscientiousness	-0.55	-0.83
Neuroticism	1.78	0.76
Openess	-0.57	4.07*
Intellect	-1.24	3.95*
Attractiveness	-8.96	0.46
Self-Esteem	-2.48*	0.24
Integrity	-3.00*	-1.34

Note. * $p < .004$.

Figure 1

Means for Self and Knowledgeable Other Reports



Note. * Self-family mean differences were statistically significant, $p < .05$.

† Self-friend mean differences were statistically significant, $p < .05$.